(4)

GL-TR-89-0100

STATISTICAL PARAMETERS FOR DESCRIBING MODEL ACCURACY

J. N. Bass
N. Grossbard
E. C. Robinson

Radex., Inc.
Three Preston Court
Bedford, MA 01730

March 20, 1989

Scientific Report No. 3

GEOPHYSICS LABORATORY
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
HANSCOM AIR FORCE BASE, MASSACHUSETTS 01731-5000

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for Public Release Distribution Unlimited
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) RX-89-0320	5. MONITORING ORGANIZATION REPORT NUMBER(S) GL-TR-89-0100
6a. NAME OF PERFORMING ORGANIZATION RADEX, Inc. 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Geophysics Laboratory
6c. ADDRESS (City, State, and ZIP Code) Three Preston Court Bedford, MA 01730	7b. ADDRESS (City, State, and ZIP Code) Hanscom AFB Massachusetts 01731-5000
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER Contract F19628-87-C-0084
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 62101F; PROJECT NO. 9993; TASK NO. XX; WORK UNIT ACCESSION NO. YN

11. TITLE (Include Security Classification)
Statistical Parameters for Describing Model Accuracy

12. PERSONAL AUTHOR(S)
* J.N. Bass, **N. Grossbard, ***E.C. Robinson

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Scientific Report #3	FROM 3/88 TO 3/89	1989 March 20	22

16. SUPPLEMENTARY NOTATION *Radex, Inc., Bedford, MA; **Boston College, Chestnut Hill, MA ***GL/LCY, Hanscom AFB, MA

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Chebychev's Inequality; Probability Density Function; Root Mean Square; Confidence Limits; Statistics; Model Accuracy

19. ABSTRACT (Continue on reverse if necessary and identify by block number)
Methods of probability theory are applied to determine parameters for describing the accuracy of a model, given a data base of measurements with which the predictions of the model are compared. From such comparisons it is customary to determine the mean and standard deviation of the measured-model difference or ratio. In the idealized case that the error distribution is Gaussian, we show that a single parameter, the root mean square (rms) of the mean and the standard deviation, approximately characterizes the accuracy of the model, since the width of the confidence interval whose center is at zero error can be represented as the product of the rms and the factor which is approximately independent of both the mean and the standard deviation. Using a modified version of Chebyshev's inequality, a similar result is obtained for the upper bound of the confidence interval width for any distribution of specified mean and standard deviation. Furthermore, for any error distribution given by a function which monotonically decreases with increasing distance from its mean, we show that the confidence that the error lies within a specified interval centered about zero decreases monotonically
(OVER)

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL E. C. Robinson	22b. TELEPHONE (Include Area Code) (617) 377-3840 — 22c. OFFICE SYMBOL AFGL/LCY

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

Abstract Continued:

with increasing magnitude of the mean. Cases in which the data set is contaminated by random and/or systematic error are incorporated into this discussion by considering distributions for the sum of two errors, or by placing an upper bound on the model error rms.

ACKNOWLEDGEMENTS

The motivation for this study arose as a result of our investigations of extensive atmospheric density data bases for F. A. Marcos of AFGL/LIS. His direction emphasized the need to evaluate the significance of various parametric measures of data-to-model ratio statistics in estimating model accuracies. We also appreciate the participation of K. H. Bhavnani of Radex, Inc. and A. Boehm of Boston College in the discussions leading to this report.

TABLE OF CONTENTS

LIST OF TABLES

1. Introduction

It is common in science and engineering to formulate mathematical models to describe complex systems. In science such models mathematically describe the behavior of systems governed by certain laws and hypotheses. A "law" in this context is some generally applicable statement, such as conservation of momentum, that is universally accepted as true. A hypothesis is an assumption, which we wish to test. A model in this case is a mathematical solution of a set of equations which represent the underlying laws and hypotheses, which are often collectively referred to as a theory. It can then be used to predict future behavior of such a system. Comparison of such predictions with observations test the validity of the model, in particular the underlying assumptions. An example is NCAR's Thermospheric General Circulation Model (TGCM)[1]. This model is a numerical solution of a set of equations expressing conservation of mass, momentum, and energy, in the thermosphere, incorporating assumptions concerning, for example, the spatial and temporal distributions of energy and momentum sources. Tests of the theory (underlying laws and assumptions) generally consists of "running" the model for a few test cases for which relevant data are available. Based on a comparison of results with the data, the theory is revised and rerun. This procedure hopefully results in improved understanding of the dynamics of the complex systems that the theory attempts to describe.

In this report, however, we will be concerned with another type of model, for example the MSIS 86 thermospheric model[2], which is used to obtain rapid, yet hopefully sufficiently accurate, estimates of key parameters which affect the design and/or operation of a man-made or natural system. These parameters may be environmental, primarily describing a system which interacts with the man-made system, or they may relate to the system itself. Thus, for example, the density of the atmosphere at the location of a satellite is an important environmental parameter, since the atmosphere exerts on the satellite a drag force which is proportional to the density. The predictions of such models are used in the design of a system, or to estimate the behavior of the system. For example, in tracking a satellite, it is usually necessary to predict its orbit over a period of time, in order to provide look directions for the tracking equipment. This requires an estimate of the atmospheric density for the prediction period.

For some such parameters a copious amount of data is available. Thus empirical models are set up with a number of adjustable parameters whose values are determined in least squares fits to the data. These models may incorporate the

dynamics of the system, as do the scientific models, but often in some grossly simplified form, to permit ease of everyday computational use. We then become concerned with the accuracy of these models in predicting the behavior they describe.

1.1 Definition of Accuracy

A dictionary definition of accuracy might be the quality of being errorless. Since such a quality cannot be realistically expected of models, we need a more practical definition. We therefore prefer the phrase "smallness of error". Thus a model is deemed accurate if the error made in using it is small. This gets us to the question of just what do we mean by the term "error".

For any single application of the model, its error may be simply stated as the difference between the true value of the quantity that the model is predicting, and the value predicted by the model. This is usually called the absolute error. In some cases, for example atmospheric density, it is often preferable to use the relative error, which is the ratio of the absolute error to the the quantity predicted by the model. Relative error is usually preferable in cases where the quantity in question varies by an order of magnitude or more over the data set, possibly causing the absolute errors to be smallest when the quantity being measured is smallest, thus giving us an unrealistic assessment of the accuracy of the model for these cases. For atmospheric density, it has been customary to use the ratio of the "true" density (as given by a measurement) to the model density, which is in fact the relative error plus 1.

The error (relative or absolute) of a model is naturally not the same for all cases. Ideally we would therefore need to specify the error for each possible instance in which the model is applied. Since this, of course, is also impossible, we therefore seek the probability density function which represents the fact that certain values of the error occur more frequently than others. The probability that a continuous variable has a value between x and x+dx is:

$$f(x)dx$$

where f(x) is thus the probability density function for the variable. The function f(x) must be positive, and satisfy

$$\int f(x)dx = 1 \qquad (1)$$

where the integration is over the entire range of x. This simply states that the probability of the variable taking on

some value within its range is unity. From the probability density function we can immediately define two commonly used measures of error, the mean error

$$\mu = \int x f(x) dx \tag{2}$$

and the standard deviation:

$$\sigma = [\int (x-\mu)^2 f(x) dx]^{1/2} \tag{3}$$

A commonly occurring probability density function is the Gaussian or normal function, which, for mean μ and standard deviation σ is given by:

$$f_g(x) = (2\pi\sigma^2)^{-1/2} \exp[-(x-\mu)^2/(2\sigma^2)] \tag{4}$$

The probability density function can in practice be estimated by comparing the model to the measurements in a properly constructed, accurate and sufficiently large data set. The data set is properly constructed if it correctly samples cases similar to those for which the model user has an interest, and omits cases that are not similar. For example, a user who needs to track a satellite which has an orbit at 1000 and 2200 hours local time is interested in the accuracy of the density model just at those two local times, not at other times. Therefore the optimum data set for this user might consist of on-board measurements by sun-synchronous satellites at 1000 and 2200.

The accuracy of the data set, the differences between measured and true values, is another matter of concern. The mean value of these errors is often called the systematic error. Often we have only a poor idea of these errors, since we cannot obtain an independent measure of them. For instance, the analysis of on-board satellite accelerometer drag measurements to yield the atmospheric density relies on an accurate value for the satellite's drag coefficient. Since the drag coefficient cannot in practice be measured under controlled conditions, it must be estimated from a model, whose accuracy in turn is poorly estimated. This could result in a roughly constant, and therefore systematic, but uncertain, error in the resulting "measured" densities.

From a given data set it is customary to compute the sample mean and standard deviation:

$$M = \Sigma x_i / N \tag{5}$$

$$s = [\Sigma (x_i - M)^2 / (N-1)]^{1/2} \tag{6}$$

as estimates of the mean μ and standard deviation σ given in Equations (2) and (3).

The data set must be sufficiently large that the sampling error in these values, which varies as $1/\sqrt{N}$, is small compared to the model errors we are determining. If the data set meets the conditions of proper sampling, negligible measurement error, and large sample size, then these estimates should be sufficiently accurate to meet practical needs. If the only problem with the data is systematic error, then the estimate s for the σ is not affected, but the estimate M for μ is.

Another way to specify accuracy is through the use of confidence limits. Given a probability density function, we can compute the total probability, the integral of the probability density function, that the error lies between two values:

$$P(a,b) = \int_a^b f(x)dx \tag{7}$$

Often we wish to find the bounds a and b of an interval whose center is at the mean of the distribution. In this report, however, we show how to define the confidence limits $\mp D$ of an interval whose center is at zero error, such that the probability $P(-\Delta,+\Delta)$ that the error lies within this interval is at least equal to some desired value. This specification has the advantage that it gives the user the desired "confidence" that the error magnitude is less than the half-width Δ of the interval.

We desire some simple parametric specification, accompanied by a table or simple computational algorithm, which would allow the user to easily derive confidence limits for any specified probability. This specification must be applicable in the presence of uncertain systematic error and non-zero mean error. We will show that if the probability density function is Gaussian, with zero or non-zero mean, then a single parameter nevertheless often suffices to derive desired confidence limits. For non-Gaussian probability density functions, a modification of Chebyshev's inequality[3] will allow us to derive from the same parameter an upper bound to the confidence interval width.

2. Methods

We discuss some simple properties of confidence limits, given a probability density function which may or may not be centered about zero. These properties are therefore useful for the case of non-zero mean error as well as zero mean. They are applicable to any probability density function, of which the Gaussian is a special case, which monotonically decreases with increasing distance from the mean.

2.1 Properties of Confidence Limits

Let $g(x)$ be a function which decreases monotonically with increasing absolute value of its argument. Therefore g is an even function of its argument x. The Gaussian function for a variable with zero mean is a special case. We define $P_A^{(g)}(L)$ as the probability that a variable whose probability density is given by g lies within the limits $\pm L-A$ ($L \geq 0$):

$$P_A^{(g)}(L) = \int_{-L-A}^{L-A} g(x)dx = \int_{-L}^{L} g(x-A)dx \tag{8}$$

The second equality expresses the fact that this is also the probability that the variable whose probability density (pd) is $g(x-A)$, with mean value A, lies within the limits $\pm L$.

We now show the following:

(i) P is an even function of A:

$$P_{-A}^{(g)} = P_A^{(g)} \tag{9}$$

Applying the change of variable $s = -x$ to the expression immediatly after the first equals sign of Eq. (8), and using the fact that g is an even function:

$$P_A^{(g)}(L) = -\int_{L+A}^{-L+A} g(s)ds$$

$$= \int_{-L+A}^{L+A} g(s)ds$$

$$= P_{-A}^{(g)}(L)$$

(ii) As a function of A, P has a single maximum at A=0:

$$P_A^{(g)}(L) < P_0^{(g)}(L) \quad (\text{if } A \neq 0) \tag{10}$$

The integral $\int_{-L+A}^{L+A} g(x)dx$ breaks down into the sum:

$$\int_{-L+A}^{-L} g(x)dx + \int_{-L}^{L} g(x)dx + \int_{L}^{L+A} g(x)dx$$

The second term is P_0. The difference P_A-P_0 is the sum of the first and third terms. We must therefore prove that this sum is negative. Since $g(x)$ is an even

function, the first term may be written as:

$$\int_{L}^{L-A} g(x)dx$$

Applying the change of variable x = L-s to the first term in the sum, and x = L+s to the other term, we obtain

$$-\int_{0}^{A} g(L-s)ds + \int_{0}^{A} g(L+s)ds$$

For A positive, these integrals are over positive s. Then the absolute value of the argument L-s of the first integrand is less than the absolute value of the argument L+s of the second integrand. Therefore the first integrand is larger than the second (both positive), and the sum is negative, as required for our proof. For negative A, the result follows from property (i) just shown above, the invariance of P_A under change of sign of A.

Finally, we note, for the Gaussian:

$$P_A^{(\sigma)}(L) = \{erf[(L+A)/(\sigma\sqrt{2})]+erf[(L-A)/(\sigma\sqrt{2})]\}/2 \qquad (11)$$

3. Results and Applications

As stated previously, sometimes we are not interested in the deviation of the model error about its mean, but rather we are interested in the total error, or put another way, the deviation of the error about zero. Thus we have to deal with the probability in a region not centered about the mean. In the previous section we gave some theorems concerning this situation, and in addition, Eq. (11) for application to a biased Gaussian, that is, a Gaussian pd with non-zero mean. We will now demonstrate, for the Gaussian, a rather surprising result that the probability for occurrence of the error within a given interval centered about zero is often approximately equal to that which would be obtained for an unbiased Gaussian pd with standard deviation equal to the root mean square (rms) of the mean and the standard deviation of the original Gaussian pd. We will next demonstrate how the rms can be used to obtain an upper bound for the confidence limits for non-Gaussian distributions. These results allow us to use the rms as a single parameter to characterize the model error in many cases.

3.1 Probabilities for a Biased Gaussian

Assume that a data point is taken from Gaussian probability density with standard deviation σ and mean (bias) $B\sigma$. Then Table 1 gives the probability for the occurrence of the value (y) within the interval

$$-\sigma\Delta x \leq y \leq \sigma\Delta x,$$

where Δx is the value indicated first column, and B is the value indicated in the top line. For example, if $\Delta x = .5$, and B = .4, then the probability of occurrence within the interval $\pm\sigma\Delta x$ is 35.58%. If $\Delta x = 1$, and B= 0, then the probability is 68.27%.

Assume that a data point is taken from a Gaussian pd with standard deviation σ and bias $B\sigma$. Then Table 2 gives the probability of occurrence of the value (y) within the interval

$$-[\sigma^2 + B^2\sigma^2]^{1/2}\Delta x_N \leq y \leq [\sigma^2 + \sigma^2B^2]^{1/2}\Delta x_N,$$

where Δx_N is the value indicated in the first column, and B is the value indicated in the top line.

Table 2 shows that by modifying σ to

$$rms = [\sigma^2 + \sigma^2B^2]^{1/2} \tag{12}$$

the probability of occurrence within symmetric intervals is approximately independent of the bias (B) when B is "small". For example the probability of occurrence of y within the interval ±.5rms is 38.29% when the bias is zero, and 38.11% when the bias is $.4\sigma$.

3.2 Probabilities for Non-Gaussian Distributions

Chebyshev's inequality states an upper bound for the probability that the value of a random variable deviates from its mean by more than a certain amount:

$$P\{|x - \mu| \geq \Delta\} \leq \sigma^2/\Delta^2$$

where μ and σ are the mean and standard deviation of the value of the random variable x, and Δ is a positive number. Brunk[3] points out that the proof of Chebyshev's inequality applies equally well to deviations of the random variable from any value, if the standard deviation σ is replaced by the square root of the mean square deviation of the variable from that value.

The proof uses a step function:

$$\theta(t) = 1, \ t \geq \Delta$$

$$= 0, \ t < \Delta$$

It is seen for $z = x - \mu$ that:

$$z^2 \geq \Delta^2\theta(|z|)$$

since, when z^2 is less then Δ^2, the right hand side is zero, and when z^2 equals or exceeds Δ^2, the right hand side is equal to Δ^2. Taking mean values:

$$\langle z^2\rangle \geq \langle\Delta^2\theta(|z|)\rangle = \Delta^2\langle\theta(|z|)\rangle = \Delta^2 P\{|z| \geq \Delta\}$$

The first equality follows from the constancy of Δ, the second from the definition of the step function. We see therefore that

$$P\{|z| \geq \Delta\} \leq \langle z^2\rangle/\Delta^2$$

Thus far no use has been made of the property that μ is the mean value of x. This is done only in the final step of the proof, substituting σ^2 for $\langle z^2\rangle$, the mean squared value of z. Thus, when considering the deviation from a value other than the mean, we may replace σ^2 in Chebyshev's inequality with the mean square deviation from the selected value. Then Chebyshev's inequality may be stated in the modified form:

$$P\{|x - x_0| \geq \Delta\} \leq \langle(x - x_0)^2\rangle/\Delta^2$$

for the deviation of x from any value x_0. For our application to the deviation of errors about zero, the mean square deviation is simply the rms defined in Eq. (12), with $\mu = B\sigma$, since

$$\langle x^2\rangle = \langle[(x - \mu) + \mu]^2\rangle$$

$$= \langle(x - \mu)^2 + 2\mu(x - \mu) + \mu^2\rangle$$

$$= \langle(x - \mu)^2\rangle + 2\mu\langle(x - \mu)\rangle + \mu^2$$

$$= \sigma^2 + \mu^2$$

We therefore finally state that

$$P\{|x| \geq \Delta\} \leq \mathrm{rms}^2/\Delta^2$$

meaning that the probability that x is outside the interval $[-\Delta, \Delta]$ cannot exceed $(\mathrm{rms}/\Delta)^2$. Thus the probability, or confidence, that the error x is within this interval is at least $1-(\mathrm{rms}/\Delta)^2$.

We may use this result to obtain an upper bound for the half width Δ of the confidence interval, such that the probability that the error x is within this interval is at least a desired value α:

$$\Delta \leq rms/(1-\alpha)^{1/2} \qquad (13)$$

We note that the upper bound given by the right hand side of Eq. (13) for 68.2% confidence is 1.75 rms, compared to 1.0 rms for Gaussians, as indicated by Table 2. For 95% confidence the upper bound given by Eq. (13) is 4.47 rms, compared to 1.96 rms for Gaussians. Futhermore we note that the interval half-width Δ from Eq. (13) can never be less than one rms.

3.3 Applications to Atmospheric Density Model Errors

AFGL has collected a substantial data base of in-situ satellite drag measurements in the lower thermosphere[4]. The densities derived from this data can then be compared with models such as the previously mentioned TGCM[1] and MSIS-86[2] model. With respect to engineering models such as the MSIS-86, the following statistics have routinely been computed:

$$M = \Sigma\, r_i/N \qquad (14)$$

$$s = [\Sigma\, (r_i-M)/(N-1)]^{1/2}/M \qquad (15)$$

where r_i is the ratio of the i^{th} measured density to the density predicted by the model for that case, and N is the number of measurements. These quantities are similar to the sample estimates, Eqs. (2) and (3), for the mean and standard deviation of the model ratio r, except that in the present case the sample estimate for the standard deviation has been divided by the mean ratio M. With these two parameters, the model user may imply that a corrected model obtained by multiplying the original model by M will have relative errors with estimated standard deviation s. There are at least two problems with this approach. One is that the user may need to know the error in the model itself, rather than in a corrected version. The second problem is that the data base may have systematic error, which introduces uncertainty in the mean M, although not in the standard deviation s. The following examples will show how to characterize the error in the model itself (not the corrected model) given that we have a sample estimate M of the mean ratio, such that a sample estimate of the mean relative error μ is M-1, where M is given by Eq. (14), and a sample estimate of the standard deviation σ given by Ms, where s is given by Eq. (15). The reason for using Ms instead of s is that we are interested here in the relative error of the original model, not the

relative error of the corrected model. The factor M is generally close to one, unless the model is very poor.

3.3.1 Model Developed from a Data Base

We assume that a model has been developed empirically from a data base. This is generally the case for models such as the MSIS-86. This data base is then used to "evaluate" the model. The mean relative model error is found to be zero, and the standard deviation is found to be 15%.

A. The data had no systematic or random error. The rms is therefore 15%. If the error distribution is Gaussian, we may then use column 1 of Table 2, to find the ± error range for any desired probability. Simply find the row containing the desired probability in column 2 and multiply the number in column 1 by 15%. If the distribution is strongly non-Gaussian, we may use Eq. (13) to find an upper bound for the error range.

B. We find out later that the measurements should have been 10% higher. Therefore, using the hypothetically corrected data base to evaluate the model obtained on the basis of the uncorrected data, we would obtain a mean error of 10% and a σ of (15 x 1.1)% = 16.5%. Thus our model error is characterized by:

$$\text{rms}(\%) = [10^2 + 16.5^2]^{1/2} = 19.3\%.$$

C. We find out later that, instead, there was uncertain systematic error in the data, in the range ± 10%. Then the total error is the sum of the measured and systematic errors. Assuming that the latter may be characterized by a probability density function with zero mean and standard deviation 10%, we obtain for the total error:

$$\text{rms} = [10^2 + 15^2]^{1/2} = 18.0\%$$

If the probability density functions of both the measured systematic errors are Gaussian, then the probability density function of the total error is also Gaussian[3], and we may use Table 2, as above, to determine the confidence limits for any specified probability. If either distribution is non-Gaussian, we may use Eq. (13).

There is another view of the systematic error which is sometimes considered: the systematic error is known definitely to be within a certain range, but its pd within that range is unknown. Thus the total error pd may be biased by some unknown amount, within this range. If the measured error distribution is Gaussian, then according to

our results of Section 2.1, property (ii), the probability P_A of occurrence within an interval centered about zero decreases with the increasing magnitude of the bias A. Therefore the worst case, within this specification, would occur if the systematic error had the maximum value, 10%. Using this worst case we get the same total rms as just given for the 10% σ uncertainty in the systematic error. The same rms is applicable if the measured error distribution is unspecified or strongly non-Gaussian, since the modified Chebyshev's inequality applied to this worst case rms produces the worst case confidence interval.

Another method of handling the systematic error, when it is known to lie within a definite finite range, is to add the maximum absolute value of the systematic error to the rms random measurement error. This is the approach suggested by Cameron[5]. It is clear that our result, using the rms of the systematic and random errors, yields tighter confidence limits.

3.3.2 Model Compared With Independent Data Base

A. The mean ratio is 1.0, the standard deviation is 15%, and there is no systematic or random measurement error. Then the rms is the same as in 3.3.1 A, 15%.

B. Same as A, except that the mean ratio is 1.1. Then, following our discussion in Sections 3.1 and 3.2,

$$\text{rms} = [10^2 + 15^2]^{1/2} = 18.0\%.$$

C. Same as B, except that the data has uncertain systematic error, and therefore an uncertainty in the mean error, within 10%. If this uncertainty is characterized by a probability density function with zero mean and standard deviation 10%, then the overall pd (measured + systematic error) is characterized by:

$$\sigma = [10^2 + 15^2]^{1/2} = 18.0\%$$

$$\mu = 10\%$$

For this biased distribution we may again define an effective rms

$$\text{rms} = [\sigma^2 + \mu^2]^{1/2} = 20.6\%$$

If if we take the second approach, mentioned above, for the systematic error, namely that we know only that it is in the interval ±10%, then, taking the worst case, as we did previously, we get a distribution whose mean is the

sum of the measured mean and the largest possible bias value:

$$\mu = 10\% + 10\% = 20\%$$

The standard deviation is as measured, 15%. Then, we obtain, for the rms:

$$rms = [20^2 + 15^2]^{1/2} = 25\%$$

D. Same as C, except that the data has random error of 5%. In this case, we know that the measured σ, estimated by s, is the standard deviation for the sum of two independent unbiased random errors: the true model (random) error (true value - model value) and the measurement error (measured value - true value). Then it is evident that the true model random error has the standard deviation

$$\sigma = [15^2 - 5^2]^{1/2} = 14.1\%.$$

Then we obtain for the rms, assuming that the systematic error uncertainty is charactized by 10% standard deviation and zero mean:

$$rms = [14.1^2 + 10^2 + 10^2]^{1/2} = 20.0\%$$

For the second view of systematic error, assuming the worst case value of 10%:

$$rms = [14.1^2 + 20^2]^{1/2} = 24.5\%$$

For Gaussian distributions we must caution that, for the conservative worst case systematic errors assumed in the last two examples, the biases obtained were larger than one sigma (1.33σ in example C, and 1.4σ in example D). For biases this large, Table 2 indicates moderate decrease in the probability of occurrence within intervals smaller than the rms, relative to the probabilities of occurrence in the same intervals for zero bias. At one rms, the probability has decreased from 68.2% to approximately 63%. However, it is seen that this trend eventually reverses as the interval increases.

4. Discussion and Conclusions

We have attempted to find a minimal set of parameters to describe model accuracy. It is hoped that, in most situations, such information would provide the potential user with all that he needs to know to assess the impact, on his situation, of any error in the model.

We have approached the problem from the standpoint of confidence limits, the bounds of an interval, centered about zero, for which the error has a desired probability of occurring. This requires determination of the probability density function for the error. We have stated two useful properties of a certain class of biased probability density functions which include the Gaussian. For the Gaussian itself we have shown that this probability often can be calculated from a single parameter, the rms of the mean and standard deviation, even when the mean is non-zero. Using a modification of Chebyshev's inequality, we have also found the minimum probability of occurrence for any distribution of specified mean and standard deviation. This leads to confidence limits considerably larger than those obtained for Gaussian distributions.

Finally we have applied these results to atmospheric density model evaluations, which are carried out by comparing the predictions of the model with measurements in a large data base. We have considered contamination of the data by random and/or systematic measurement errors. Although the relevant error sources are often Gaussian, this is not always the case. Marcos, et. al., have found occasional departures from the Gaussian in their evaluations[6]. In addition to the mean and standard deviation, two additional parameters, called skewness and kurtosis, were used by them to characterize these cases. If the skewness is zero, then the function is still symmetric about its mean, so that we may still be able to apply the two properties mentioned in Section 2. This would not be the case if the skewness is non-zero, for this would indicate that the function is not symmetric about its mean. Such a function cannot therefore depend only on the absolute value of the difference between its argument and the mean, and therefore cannot fall into the class of functions discussed there. However if these departures are small, and we need to decide just what "small" is, these results may still be applicable. Although Chebyshev's inequality could be applied to these cases, it is evident that confidence limits considerably closer to those for Gaussian should be obtainable. Therefore we should look for methods to extend to near-Gaussians the results obtained here for Gaussians.

References

1. Dickinson, R. E., E. C. Ridley, and R. G. Roble, "A Three Dimensional General Circulation Model of the Thermosphere", J. Geophys. Res., Vol. 86, Pp. 1499-1512, 1981.

2. Hedin, A. E., "MSIS-86 Thermospheric Model", J. Geophys. Res., Vol. 92, No. A5, Pp. 4649-4662, 1987.

3. Brunk, H. D., An Introduction to Mathematical Statistics, Xerox College Publishing, Lexington, MA/Toronto, 1975.

4. Marcos, F. A., "Requirements for Improved Thermospheric Neutral Density Models, AAS/AIAA Astrodynamics Specialist Conference", Paper No. AAS 85-312, 1985.

5. Cameron, J. M., "Chapter 2. Statistics", in Fundamental Formulas of Physics, Vol. 1, D. H. Menzel, ed., Dover Publications, Inc., New York, 1960.

6. Marcos, F. A., R. E. McInerney, and R. W. Fioretti, "Variability of the Lower Thermosphere Determined from Satellite Accelerometer Data", Air Force Geophysics Laboratory, Hanscom AFB, MA, AFGL-TR-78-0134, ADA 058982, 1978.

TABLE SHOWING PER-CENT OF READINGS TAKEN FROM A GAUSSIAN DISRTIBUTION WITH SIGMA=1 AND THE INDICATED BIASES

ERROR-BAR -TO+ BIAS=	.00	.20	.40	.60	.80	1.00	1.20	1.40	1.60	1.80	2.00
.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
.10	7.97	7.81	7.36	6.66	5.79	4.84	3.89	3.00	2.22	1.58	1.09
.20	15.85	15.54	14.65	13.27	11.56	9.68	7.79	6.03	4.48	3.20	2.20
.30	23.58	23.13	21.82	19.80	17.29	14.52	11.73	9.11	6.81	4.89	3.38
.40	31.08	30.50	28.81	26.21	22.95	19.35	15.71	12.27	9.23	6.69	4.66
.50	38.29	37.59	35.58	32.45	28.53	24.17	19.74	15.53	11.78	8.61	6.06
.60	45.15	44.36	42.06	38.49	34.00	28.98	23.83	18.91	14.48	10.69	7.61
.70	51.61	50.74	48.22	44.30	39.34	33.75	27.98	22.41	17.33	12.95	9.33
.80	57.63	56.71	54.04	49.85	44.52	38.48	32.18	26.03	20.37	15.40	11.25
.90	63.19	62.24	59.47	55.11	49.53	43.15	36.42	29.78	23.58	18.06	13.38
1.00	68.27	67.31	64.50	60.06	54.33	47.72	40.68	33.64	26.96	20.93	15.73
1.10	72.87	71.91	69.12	64.69	58.92	52.20	44.94	37.59	30.51	24.01	18.31
1.20	76.99	76.06	73.33	68.98	63.27	56.54	49.18	41.61	34.20	27.29	21.12
1.30	80.64	79.75	77.14	72.93	67.36	60.72	53.36	45.67	38.02	30.76	24.15
1.40	83.85	83.01	80.54	76.54	71.18	64.72	57.46	49.74	41.94	34.39	27.39
1.50	86.64	85.86	83.56	79.81	74.73	68.53	61.44	53.80	45.92	38.16	30.83
1.60	89.04	88.33	86.22	82.74	77.99	72.11	65.29	57.79	49.93	42.04	34.44
1.70	91.09	90.45	88.53	85.36	80.97	75.46	68.96	61.69	53.93	45.99	38.20
1.80	92.81	92.25	90.53	87.67	83.67	78.56	72.44	65.47	57.89	49.98	42.07
1.90	94.26	93.76	92.25	89.70	86.09	81.41	75.71	69.10	61.77	53.97	46.01
2.00	95.45	95.02	93.70	91.46	88.24	84.00	78.75	72.54	65.53	57.92	50.00
2.10	96.43	96.06	94.92	92.97	90.13	86.34	81.55	75.78	69.14	61.79	53.98
2.20	97.22	96.91	95.94	94.26	91.79	88.42	84.10	78.80	72.57	65.54	57.92
2.30	97.86	97.59	96.78	95.36	93.22	90.27	86.41	81.58	75.80	69.14	61.79
2.40	98.36	98.14	97.47	96.27	94.45	91.89	88.48	84.13	78.81	72.57	65.54
2.50	98.76	98.58	98.03	97.03	95.50	93.30	90.31	86.43	81.59	75.80	69.15
2.60	99.07	98.92	98.47	97.66	96.37	94.50	91.92	88.49	84.13	78.81	72.57
2.70	99.31	99.19	98.83	98.17	97.11	95.53	93.31	90.32	86.43	81.59	75.80
2.80	99.49	99.40	99.11	98.58	97.71	96.40	94.52	91.92	88.49	84.13	78.81
2.90	99.63	99.56	99.33	98.90	98.20	97.12	95.54	93.32	90.32	86.43	81.59
3.00	99.73	99.68	99.50	99.16	98.60	97.72	96.41	94.52	91.92	88.49	84.13
3.10	99.81	99.77	99.63	99.37	98.92	98.21	97.13	95.54	93.32	90.32	86.43
3.20	99.86	99.83	99.73	99.53	99.18	98.61	97.72	96.41	94.52	91.92	88.49
3.30	99.90	99.88	99.80	99.65	99.38	98.93	98.21	97.13	95.54	93.32	90.32
3.40	99.93	99.92	99.86	99.74	99.53	99.18	98.61	97.72	96.41	94.52	91.92
3.50	99.95	99.94	99.90	99.81	99.65	99.38	98.93	98.21	97.13	95.54	93.32
3.60	99.97	99.96	99.93	99.86	99.74	99.53	99.18	98.61	97.72	96.41	94.52
3.70	99.98	99.97	99.95	99.90	99.81	99.65	99.38	98.93	98.21	97.13	95.54
3.80	99.99	99.93	99.96	99.93	99.86	99.74	99.53	99.18	98.61	97.72	96.41
3.90	99.99	99.99	99.98	99.95	99.90	99.81	99.65	99.38	98.93	98.21	97.13
4.00	99.99	99.99	99.98	99.97	99.93	99.86	99.74	99.53	99.18	98.61	97.72
4.10	100.00	99.99	99.99	99.98	99.95	99.90	99.81	99.65	99.38	98.93	98.21
4.20	100.00	100.00	99.99	99.98	99.97	99.93	99.87	99.74	99.53	99.18	98.61
4.30	100.00	100.00	100.00	99.99	99.98	99.95	99.90	99.81	99.65	99.38	98.93
4.40	100.00	100.00	100.00	99.99	99.98	99.97	99.93	99.87	99.74	99.53	99.18
4.50	100.00	100.00	100.00	100.00	99.99	99.98	99.95	99.90	99.81	99.65	99.38
4.60	100.00	100.00	100.00	100.00	99.99	99.98	99.97	99.93	99.87	99.74	99.53
4.70	100.00	100.00	100.00	100.00	100.00	99.99	99.98	99.95	99.90	99.81	99.65
4.80	100.00	100.00	100.00	100.00	100.00	99.99	99.98	99.97	99.93	99.87	99.74
4.90	100.00	100.00	100.00	100.00	100.00	100.00	99.99	99.98	99.95	99.90	99.81
5.00	100.00	100.00	100.00	100.00	100.00	100.00	99.99	99.98	99.97	99.93	99.87

Table 1: Probabilities for a Biased Gaussian vs. Fraction of Standard Deviation

TABLE SHOWING PER-CENT OF READINGS TAKEN FROM A GAUSSIAN DISRTIBUTION WITH SIGMA=1 AND THE INDICATED BIASES

ERROR-BAR -TO+ FRACTION OF R.M.S. / BIAS=	.00	.20	.40	.60	.80	1.00	1.20	1.40	1.60	1.80	2.00
.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
.10	7.97	7.96	7.92	7.76	7.41	6.84	6.08	5.18	4.22	3.30	2.47
.20	15.85	15.85	15.76	15.45	14.78	13.69	12.22	10.49	8.67	6.91	5.30
.30	23.58	23.57	23.45	23.01	22.06	20.52	18.47	16.07	13.54	11.07	8.81
.40	31.08	31.07	30.93	30.38	29.21	27.33	24.86	21.99	18.97	16.01	13.26
.50	38.29	38.28	38.11	37.49	36.17	34.09	31.38	28.28	25.02	21.82	18.80
.60	45.15	45.14	44.95	44.29	42.90	40.75	38.01	34.90	31.68	28.50	25.47
.70	51.61	51.59	51.40	50.73	49.35	47.27	44.67	41.78	38.83	35.93	33.17
.80	57.63	57.61	57.43	56.78	55.48	53.57	51.26	48.78	46.30	43.90	41.63
.90	63.19	63.17	63.00	62.40	61.24	59.60	57.70	55.74	53.86	52.11	50.49
1.00	68.27	68.26	68.10	67.57	66.59	65.28	63.85	62.48	61.26	60.22	59.33
1.10	72.87	72.86	72.72	72.26	71.50	70.55	69.61	68.83	68.27	67.90	67.71
1.20	76.99	76.98	76.86	76.52	75.96	75.36	74.89	74.66	74.66	74.89	75.28
1.30	80.64	80.63	80.55	80.30	79.96	79.69	79.63	79.85	80.31	80.97	81.78
1.40	83.85	83.84	83.78	83.63	83.49	83.50	83.78	84.34	85.12	86.05	87.09
1.50	86.64	86.64	86.60	86.54	86.56	86.80	87.33	88.11	89.07	90.12	91.21
1.60	89.04	89.04	89.03	89.04	89.20	89.61	90.30	91.19	92.20	93.25	94.27
1.70	91.09	91.09	91.09	91.17	91.43	91.95	92.72	93.63	94.60	95.55	96.42
1.80	92.81	92.82	92.84	92.96	93.29	93.87	94.65	95.51	96.38	97.17	97.86
1.90	94.26	94.26	94.29	94.45	94.82	95.41	96.14	96.92	97.64	98.27	98.77
2.00	95.45	95.45	95.50	95.67	96.05	96.62	97.28	97.94	98.51	98.98	99.33
2.10	96.43	96.43	96.48	96.66	97.03	97.55	98.12	98.66	99.09	99.42	99.65
2.20	97.22	97.22	97.27	97.46	97.80	98.26	98.73	99.15	99.46	99.68	99.82
2.30	97.86	97.86	97.91	98.08	98.40	98.79	99.16	99.47	99.69	99.83	99.92
2.40	98.36	98.36	98.41	98.57	98.84	99.17	99.46	99.68	99.83	99.92	99.96
2.50	98.76	98.76	98.81	98.95	99.18	99.44	99.66	99.81	99.91	99.96	99.98
2.60	99.07	99.07	99.11	99.24	99.43	99.63	99.79	99.89	99.95	99.98	99.99
2.70	99.31	99.31	99.35	99.45	99.61	99.76	99.87	99.94	99.96	99.99	100.00
2.80	99.49	99.49	99.52	99.61	99.73	99.85	99.92	99.97	99.99	100.00	100.00
2.90	99.63	99.63	99.66	99.73	99.82	99.90	99.96	99.98	99.99	100.00	100.00
3.00	99.73	99.73	99.75	99.81	99.88	99.94	99.98	99.99	100.00	100.00	100.00
3.10	99.81	99.81	99.83	99.87	99.92	99.96	99.99	100.00	100.00	100.00	100.00
3.20	99.86	99.86	99.88	99.91	99.95	99.98	99.99	100.00	100.00	100.00	100.00
3.30	99.90	99.90	99.92	99.94	99.97	99.99	100.00	100.00	100.00	100.00	100.00
3.40	99.93	99.93	99.94	99.96	99.98	99.99	100.00	100.00	100.00	100.00	100.00
3.50	99.95	99.95	99.96	99.97	99.99	100.00	100.00	100.00	100.00	100.00	100.00
3.60	99.97	99.97	99.97	99.98	99.99	100.00	100.00	100.00	100.00	100.00	100.00
3.70	99.98	99.98	99.98	93.99	100.00	100.00	100.00	100.00	100.00	100.00	100.00
3.80	99.99	99.99	99.99	99.99	100.00	100.00	100.00	100.00	100.00	100.00	100.00
3.90	99.99	99.99	99.99	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
4.00	99.99	99.99	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
4.10	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
4.20	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
4.30	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
4.40	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
4.50	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
4.60	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
4.70	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
4.80	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
4.90	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
5.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Table 2: Probabilities for a Biased Gaussian vs. Fraction of Root Mean Square